RAW SEQUENCE LISTING ERROR REPORT

BIOTECHNOLOGY SYSTEMS BRANCH
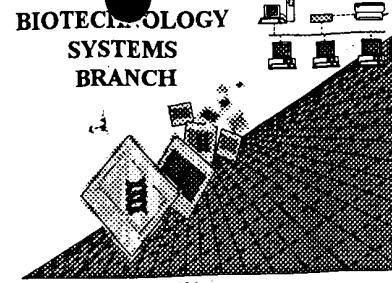
The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _09/939,408_

Source: _OIPE_

Date Processed by STIC: _9/18/2001_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov **or phone 703-306-4119 (R. Wax)**
PATENTIN 3.0 e-mail help: patin3help@uspto.gov **or phone 703-306-4119 (R. Wax)**

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE <u>CHECKER VERSION 3.0 PROGRAM</u>, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).
Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

<u>Checker Version 3.0 can be down loaded from the USPTO website at the following address:</u>
http://www.uspto.gov/web/offices/pac/checker

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: 09/939,408

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 ____ Wrapped Nucleics
Wrapped Aminos

The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2 ✓ ____ Invalid Line Length

The rules require that a line not exceed 72 characters in length. This includes white spaces.

3 ____ Misaligned Amino
Numbering

The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4 ____ Non-ASCII

The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5 ____ Variable Length

Sequence(s) ____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6 ____ PatentIn 2.0
"bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) ____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

7 ____ Skipped Sequences
(OLD RULES)

Sequence(s) ____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i)   SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8 ____ Skipped Sequences
(NEW RULES)

Sequence(s) ____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9 ____ Use of n's or Xaa's
(NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10 ____ Invalid <213>
Response

Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

11 ____ Use of <220>

Sequence(s) ____ missing the <220> "Feature" and associated numeric identifiers and responses.
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32)   (Sec. 1.823 of Sequence Rules)

12 ____ PatentIn 2.0
"bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13 ____ Misuse of n

n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

OIPE

RAW SEQUENCE LISTING   DATE: 09/18/2001
PATENT APPLICATION: US/09/939,408   TIME: 11:04:26

Input Set : A:\29479-500nsca sequence listing.txt
Output Set: N:\CRF3\09182001\I939408.raw

```
  3 ·110:> APPLICANT: Yoshida, Roberta
  4          Kootstra, Anna
  7 ·120:> TITLE OF INVENTION: Phenylalanine Ammonia Lyase Polypeptide and
  8          Polynucleotide Sequences and Methods of Obtaining and
  9          Using Same
 11 <130:> FILE REFERENCE: 29479/500NSCA
 14 <140> CURRENT APPLICATION NUMBER: US/09/939,408
 14 <141> CURRENT FILING DATE: 2001-08-24
 14 <150:> PRIOR APPLICATION NUMBER: US 09/624,693
 17 ·151:> PRIOR FILING DATE: 2000-07-24
 18 ·150:> PRIOR APPLICATION NUMBER: PCT/US01/23270
 21 ·151:> PRIOR FILING DATE: 2001-07-24
 24 <160> NUMBER OF SEQ ID NOS: 30  OK
 26 ·170:> SOFTWARE: PatentIn Ver. 2.0
```

E-->

ERRORED SEQUENCES

see p. 2 and p. 3

FYI

```
<210> 20
<211> 2475
<212> DNA
<213> Artificial Sequence


<220>
<221> misc_difference
<222> (13, 34, 46, 49, 51, 57, 59, 68, 69, 73, 75 - 77, 79, 82, 84,
102, 106, 108, 115, 119, 144, 145, 164, 168, 190, 191, 195, 199 -      201,
233, 251, 264, 266, 276, 284, 312, 315, 330, 331, 333, 340,       348, 357,
405, 423, 450, 456, 463, 468, 475, 483, 555, 570, 675,       681, 715, 716,
723, 729, 730, 732, 743, 744, 758, 783, 921, 963,       1042, 1043, 1176,
1197, 1241, 1270, 1281, 1308, 1380, 1383, 1407,       1446, 1449, 1452, 1470,
1488, 1509, 1542, 1554, 1563, 1617, 1675,       1677, 1678, 1681, 1683, 1684,
1690, 1693, 1708, 1710, 1723 -       1735, 1745, 1762, 1768, 1776, 1855 -
1862, 1872 - 1875, 1880,       1881, 1895, 1950, 1952, 1962, 1971, 1976, 2001,
2145, 2146, 2151,       2183, 2187 - 2189, 2191, 2193 - 2195, 2197, 2199,
2206, 2208 -       2212, 2215 - 2217, 2219, 2221 - 2223, 2226 - 2233, 2236,
2239,       2241 -2243, 2247, 2248, 2251, 2254, 2256 - 2260, 2265, 2266,
2268, 2269, 2271, 2272, 2274 - 2280, 2282 - 2285, 2287, 2289,       2290,
2293, 2294, 2298, 2300 - 2303, 2305, 2307, 2308, 2312,       2313, 2315 -
2319, 2322 - 2324, 2326, 2327, 2329 - 2335, 2337 -       2339, 2341, 2344 -
2346, 2349 - 2351, 2354, 2356, 2358 - 2363,       2365, 2366, 2368, 2371,
2373, 2374, 2377, 2379, 2380, 2382, 2384,       2385, 2387, 2390, 2392, 2393,
2395 - 2403, 2405, 2409 - 2475)
<223> n = A or C or G or T; "n" indicates no consensus at that position
<223> Description of Artificial Sequence: Consensus
       Sequence of SEQ ID NOs: 12, 16, and 18
```

(See, also, item 2 on Error Summary sheet
Please ensure all lines NOT exceed 72 characters.)

See 1.823 d) Sequence Rules

This response needs to be shortened; its size is causing the <160> error on p. 1.

Please list just the number (location) of the first and the last bases e.g. (13)...(2475)

do same in Sequence 21, too

<400> 28

```
atg gcm cct tcc ttg gac tcg ctc gcc acc acg ctc gcc aac ggc ttt      48
Met Ala Pro Ser Leu Asp Ser Leu Ala Thr Thr Leu Ala Asn Gly Phe
1               5                   10                  15
acc aac ggc tcg cac gcc gct ccg acc aag tcg gct gcg ggc ccc act ─)    96
Thr Asn Gly Ser His Ala Ala Pro Thr Lys Ser Ala Ala Gly Pro Thr
            20                  25                  30
```

move over

codons must
be directly
above respective
amino acids

Input Set : **A:\29479-500nsca sequence listing.txt**
Output Set: **N:\CRF3\09182001\I939408.raw**

L:14 M:270 C: Current Application Number differs, Replaced Current Application No
L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:56 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:74 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:192 M:283 W: Missing Blank Line separator, <220> field identifier
L:245 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:408 M:283 W: Missing Blank Line separator, <400> field identifier
L:436 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:24 M:203 E: No. of Seq. differs, <160> Number Of Sequences:Input (30) Counted (19)